AP Stat Worksheet

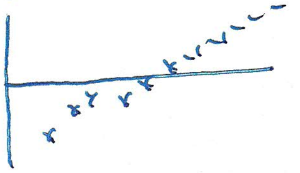
① Chips Ahoy Problem.

In order to to find the 95% confidence interval, I will need to satisfy the Independence Assumption & The Normal Population Assumption. In order to satisfy the Independence Assumption, I will meet the

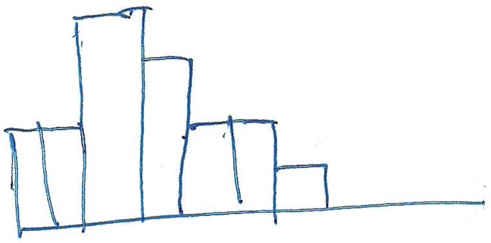
Randomization Condition: The bags were selected at random.

10% Condition: I will assume that the number of bags in the sample is less than 10% of all bags.

To satisfy the Normal Population Assumption, I will meet the Nearly Normal Assumption:



The normal probability plot is nearly straight indicating that the distribution is approximately symmetric



The histogram shows that the distribution is unimodal

Given the sample size, an approximately symmetric, unimodal distribution will satisfy the Condition.

Since I have met all the Assumptions/Conditions, I can use a t-model with 15 degrees of freedom to construct an 95% confidence interval.

$\bar{x} = 1238.0625$

$S_x = 94.31$

t^*_{15} (for 95% CI) $= 2.13$

$SE(\bar{x}) = \frac{S_x}{\sqrt{n}} = \frac{94.31}{\sqrt{16}}$

$= 23.58$

$\bar{x} \pm ME$

$\bar{x} \pm t^*_{15} SE(\bar{x})$

$1238.0625 \pm 2.13(23.58)$

$(1187.8, 1288.3)$

We are 95% confident that the true mean of choc. chips in a bag of Chip Ahoy is between 1187.8 & 1288.3 chips.

AP Worksheet Solutions

~~1. Chips Ahoy Problem~~

2. Hours of TV watched Problem

$$H_0 : \mu = 13$$
$$H_A : \mu > 13$$

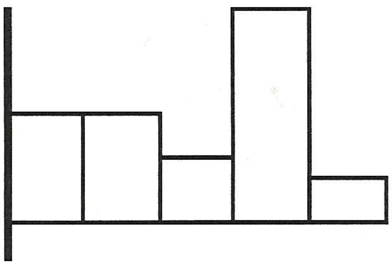
To conduct a one-sample *t*-test, I need to satisfy the Independence Assumption and the Normal Population Assumption. To meet the Independence Assumption, I will satisfy the

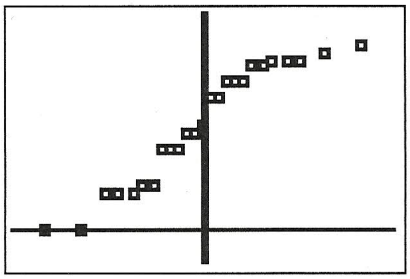
Randomization Condition: The students were picked randomly

10% Condition: We can assume that the 25 students are less than 10% of the school's population.

To meet the Normal Probability Assumption, I will satisfy the

Nearly Normal Condition:





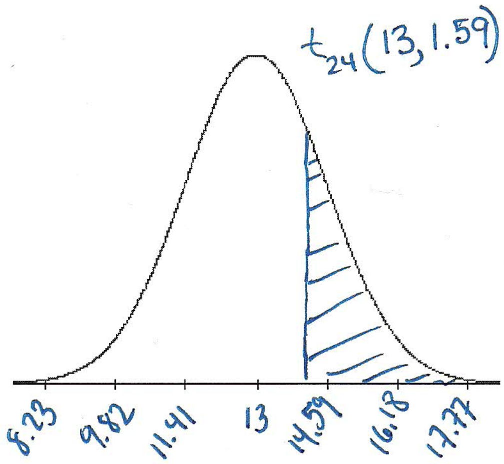
The sample distribution is unimodal (as can be seen the in the histogram) and is fairly symmetric (as can be seen in both the histogram and the normal probability plot). Given the moderate sample size of 25, this is adequate to satisfy the condition.

Since I have satisfied the Assumptions/Conditions, I can now use the *t*-distribution with 24 degrees of freedom to model the sampling distribution of sample means. I can conduct a one sample *t*-test.

$\bar{x} = 14.32$

$S_x = 7.96$

$$SE(\bar{x}) = \frac{S_x}{\sqrt{n}} = \frac{7.96}{\sqrt{25}} = 1.59$$



$$t_{24} = \frac{\bar{x} - \mu}{SE(\bar{x})} = \frac{14.32 - 13}{1.59} = .83$$

$$\text{P-value} = p(\bar{x} > 14.32)$$
$$= p(z > .83) = .208$$

Given a P-value of .208, I can expect results like this (or greater) 208 out of 1000 times. This is not unusual and, as a result, I fail to reject the null hypothesis at an alpha level of 0.05. The evidence is not strong enough to reject the claim made that student's average 13 hours of watching TV in a week. Our results were not statistically significant.

Calculator Output:

```
T-Test
 μ>13
 t=.8289381764
 p=.2076539258
 x̄=14.32
 Sx=7.96199305
 n=25
```

3. Fat content in Pizza brands Problem

$$H_0 : \mu_D - \mu_{PJ} = 0$$
$$H_A : \mu_D - \mu_{PJ} \neq 0$$

To conduct a two-sample *t*-test, I need to satisfy the Independence Assumption and the Normal Population Assumption. To meet the Independence Assumption, I will satisfy the

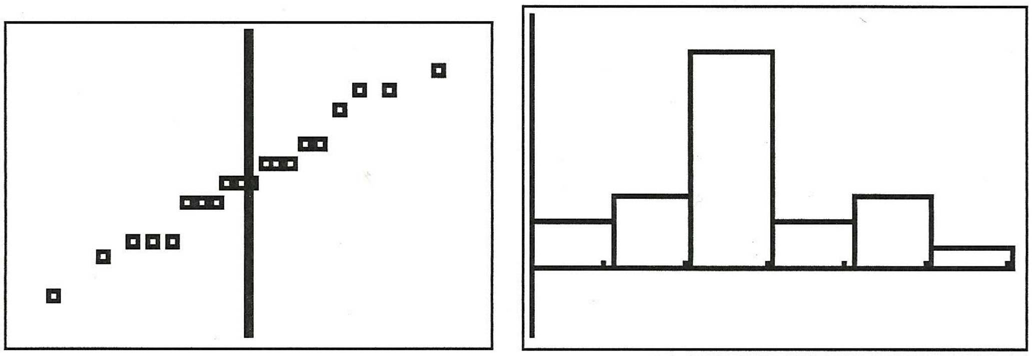
Randomization Condition: The pizzas were selected randomly in both samples

10% Condition: We can assume that the 20 and 15 pizza from each brand are less than 10% of the pizza made by each company

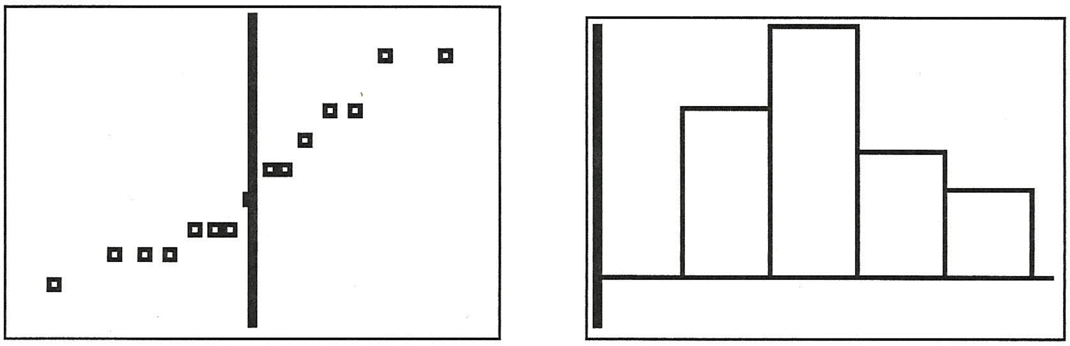
To meet the Normal Probability Assumption, I will satisfy the

Nearly Normal Condition:

Brand D



Brand PJ



Both sample distributions are unimodal and both are relatively symmetric. Given the moderate sample sizes, this is adequate to satisfy the condition.

Since I have satisfied the Assumptions/Conditions, I can now use the *t*-distribution with 32.8 degrees of freedom to model the sampling distribution of sample means. I can conduct a two sample *t*-test.

$\bar{x}_D = 11.25$ $\qquad \bar{x}_{PJ} = 6.53$

$S_{x_D} = 3.19$ $\qquad S_{x_{PJ}} = 2.59$

$n_D = 20$ $\qquad n_{PJ} = 15$

$\bar{x}_D - \bar{x}_{PJ} = 4.72$

$SE(\bar{x}_D - \bar{x}_{PJ}) = \sqrt{\dfrac{S_D^2}{n_D} + \dfrac{S_{PJ}^2}{n_{PJ}}} = \sqrt{\dfrac{(3.19)^2}{20} + \dfrac{(2.59)^2}{15}} = .98$

P-value $=$

$= 2p(\bar{x}_D - \bar{x}_{PJ} > 4.72)$

$= 2p(t_{24} > 4.82)$

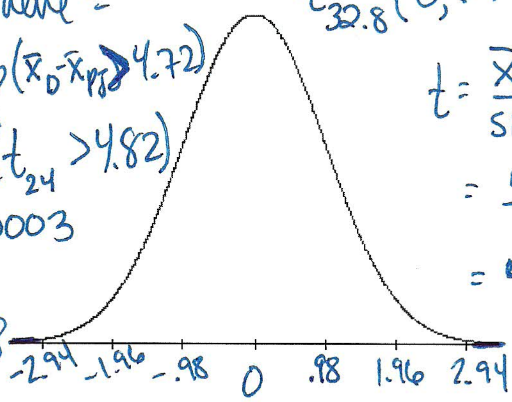
$= .00003$

$t_{32.8}(0, .98)$ $\quad \bar{x}_D - \bar{x}_{PJ}$

$t = \dfrac{\bar{x} - \mu}{SE(\bar{x}_D - \bar{x}_{PJ})}$

$= \dfrac{4.72 - 0}{.98}$

$= 4.82$



Given a P-value of .00003, we would expect that if the null was true that we would have a mean this different or more about 3 out of 100,000 times. That P-Value is small enough (at an alpha level of 0.05) that I reject the null hypothesis that the fat content of the two brands of pizza is the same. There appears to be enough evidence to say that there is a difference in the amount of fat in the two brands of pizza. Our results are statistically significant.

Calculator Output:

```
2-SampTTest
 μ1≠μ2
 t=4.823439243
 p=3.1511544E-5
 df=32.75726842
 x̄1=11.25
↓x̄2=6.533333333
```